

Influence of Expected Reward on Temporal Order Judgment

Mohsen Rakhshan^{1*}, Vivian Lee^{1*}, Emily Chu^{1*}, Lauren Harris²,
Lillian Laiks³, Peyman Khorsand¹, and Alireza Soltani¹

Abstract

■ Perceptual decision-making has been shown to be influenced by reward expected from alternative options or actions, but the underlying neural mechanisms are currently unknown. More specifically, it is debated whether reward effects are mediated through changes in sensory processing, later stages of decision-making, or both. To address this question, we conducted two experiments in which human participants made saccades to what they perceived to be either the first or second of two visually identical but asynchronously presented targets while we manipulated expected reward from correct and incorrect responses on each trial. By comparing reward-induced bias in target selection (i.e., reward bias) during the two experiments, we determined whether reward caused changes in sensory or decision-making processes. We found similar reward biases in the two experiments indicating that reward informa-

tion mainly influenced later stages of decision-making [R1.1]. Moreover, the observed reward biases were independent of the individual's sensitivity to sensory signals. This suggests that reward effects were determined heuristically via modulation of decision-making processes instead of sensory processing. To further explain our findings and uncover plausible neural mechanisms, we simulated our experiments with a cortical network model and tested alternative mechanisms for how reward could exert its influence. We found that our experimental observations are more compatible with reward-dependent input to the output layer of the decision circuit. Together, our results suggest that, during a temporal judgment task, reward exerts its influence via changing later stages of decision-making (i.e., response bias) rather than early sensory processing (i.e., perceptual bias). ■

INTRODUCTION

Imagine deciding whether your iPhone or your friend's Google Pixel takes sharper photos. To make this decision impartially, you look at some photos on both phones but end up favoring your own phone. Your decision could be impartial and entirely based on perceived quality of photos on the two phones but could also be influenced by fun memories of taking photos with your phone. Similarly, any perceptual decision-making could depend not only on sensory evidence but also on experienced or expected reward (Sugrue, Corrado, & Newsome, 2005). Understanding how reward information is incorporated into perceptual choice can provide valuable insights into the neural mechanisms underlying both decision-making and reward processes (Farashahi, Ting, Kao, Wu, & Soltani, 2018; Christopoulos, Bonaiuto, & Andersen, 2015; Christopoulos & Schrater, 2015; Gao, Tortell, & McClelland, 2011; Rorie, Gao, McClelland, & Newsome, 2010; Stanford, Shankar, Massoglia, Costello, & Salinas, 2010; Sugrue et al., 2005).

Recently, there have been a number of studies that investigated mechanisms by which reward information influences perceptual decision-making (Farashahi, Ting, et al., 2018; Rajsic, Perera, & Pratt, 2017; Tosoni, Committeri, Calluso, & Galati, 2017; Cicmil, Cumming, Parker, & Krug, 2015; Gao et al., 2011; Diederich, 2008; Liston & Stone, 2008; Voss, Rothermund, & Brandtstädter, 2008; Diederich & Busemeyer, 2006). Some of these studies suggest that reward information mainly affects perceptual choice by altering the starting or end point of decision-making processes (Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012; Gao et al., 2011; Rorie et al., 2010; Summerfield & Koechlin, 2010; Feng, Holmes, Rorie, & Newsome, 2009; Diederich, 2008; Diederich & Busemeyer, 2006). For example, Diederich and colleagues propose that value-based perceptual decision-making follows a two-stage process in which the payoff of the alternative choices is evaluated first without directly influencing the processing of sensory information that happens later (Diederich, 2008; Diederich & Busemeyer, 2006). However, others argue that reward directly influences the processing of sensory information and perception (Cicmil et al., 2015; Liston & Stone, 2008; Pleger, Blankenburg, Ruff, Driver, & Dolan, 2008; Voss et al., 2008).

¹Dartmouth College, ²Florida State University, ³Texas A&M University

*Equal contribution.

These competing hypotheses—influence of reward on sensory processing or influence on later stages of decision-making—predict that unequal expected reward should result in perceptual bias or response bias, respectively (Liston & Stone, 2008). That is, by changing sensory processing, reward could alter perception (perceptual bias) similarly to the effect of selective attention on contrast judgment (Carrasco & Barbot, 2019; Carrasco, Ling, & Read, 2004). Going back to our phone analogy, reward could make the more rewarding photos to appear sharper. On the other hand, modulation of later decision processes could bias response toward the more rewarding option (response bias) without any changes in perception. In our analogy, this corresponds to favoring our phone without perceiving any difference in image quality. Nonetheless, evidence supporting either hypothesis mostly has been based on fitting performance and RT data using different models (e.g., drift diffusion model) and thus is model dependent. Moreover, it is unclear whether positive and negative expected reward outcomes influence perceptual decision-making similarly or differently.

To address these questions and distinguish between the two alternative hypotheses, we used two sets of experiments to directly measure the influence of unequal expected reward on perceptual decision-making during a temporal judgment task. Our design resembles that of a study by Shore and colleagues (Shore, Spence, & Klein, 2001) in which the authors used attentional cueing in a temporal order judgment (TOJ) task and asked participants to report the first or second targets that appeared on the screen. By comparing the shifts in psychometric functions in the two tasks, Shore and colleagues aimed to separate changes in sensory processing from response biases [R1.5]. Similarly, in our experiments, participants made saccades to report what they perceived to be the first (Experiment 1) or second (Experiment 2) of the two identical targets that appeared on the computer screen with varying onset asynchrony. Instead of attentional cueing, however, we manipulated the amount of reward points to be gained (gains) or lost (losses) upon correct and incorrect response, respectively, on each trial. Different values of expected gains and losses associated with the left and right choices were presented on the two sides of the fixation cross to create three reward conditions: Neutral, Gain, and Loss. This design allowed us to estimate the shift in target selection and sensitivity of choice in response to different target onset asynchronies (TOAs) and how this shift and sensitivity were affected by reward manipulation.

In both our experiments, the expected reward associated with left and right choices did not predict the correct response on a given trial. As a result, the observer or decision-maker could either ignore reward information or use this information to enhance sensory processing on both sides equally. However, none of these observers (which we refer to as Type 1 observers) would exhibit bias in target selection [R1.2] [R1.4].

Alternatively, the observer could utilize differential reward information to guide sensory processing and target selection via several different mechanisms. First, the observer could use reward information to attend to the more rewarding side more strongly (Type 2 observers), which could subsequently result in enhanced processing of visual information on that side. Such an enhancement of visual processing would cause the target on the more rewarding side (better target) to appear earlier even if the both targets appeared on the screen simultaneously (perceptual bias). This would, in turn, increase the probability of choosing the better target in Experiment 1 (Figure 1A). In contrast, the earlier perception of the better target would increase the probability of choosing the worse target in Experiment 2 in which the participant has to saccade to the target that appeared second (Figure 1B). Second, the observer could use reward information to directly bias their response to increase their overall payoff. Such response bias could increase the overall payoff because there are trials in which the temporal judgment is very difficult (for small TOA) or impossible (when TOA was equal to 0) and it is beneficial to choose the better target in those trials. This effect of reward on decision-making results in more frequent selection of the better target in both Experiments 1 and 2 (Figure 1C–D). The response bias could be dependent (Type 3 observers) or independent (Type 4 observers) of the individual observer's sensitivity to visual information; however, only the former observers are able to optimize the overall payoff [R1.2] [R1.4].

Therefore, changes in sensory processing would result in perceptual bias and opposite shifts in target selection in Experiments 1 and 2 for Type 2 observers, but changes in later stages of decision-making would cause response bias and similar shifts in target selection toward the better target in the two experiments (Types 3 and 4 observers). Moreover, these mechanisms predict different relationships between the shift in target selection due to reward and the overall sensitivity to sensory information (see Results for more details) [R1.2] [R1.4].

Therefore, by comparing reward-induced shifts in target selection in the two experiments and examining the relationship between these shifts and sensitivity to sensory information, we aimed to identify mechanisms by which reward exerts its influence on perceptual decision-making [R1.2]. We also used a biophysically plausible cortical network model to replicate the experimental data to identify possible neural mechanisms underlying the influence of reward on perceptual choice.

METHODS

Ethics Statement

A total of 29 (15 female) participants were recruited from the Dartmouth College student population (ages 18–22 years) to participate in our experiments. Of the 29

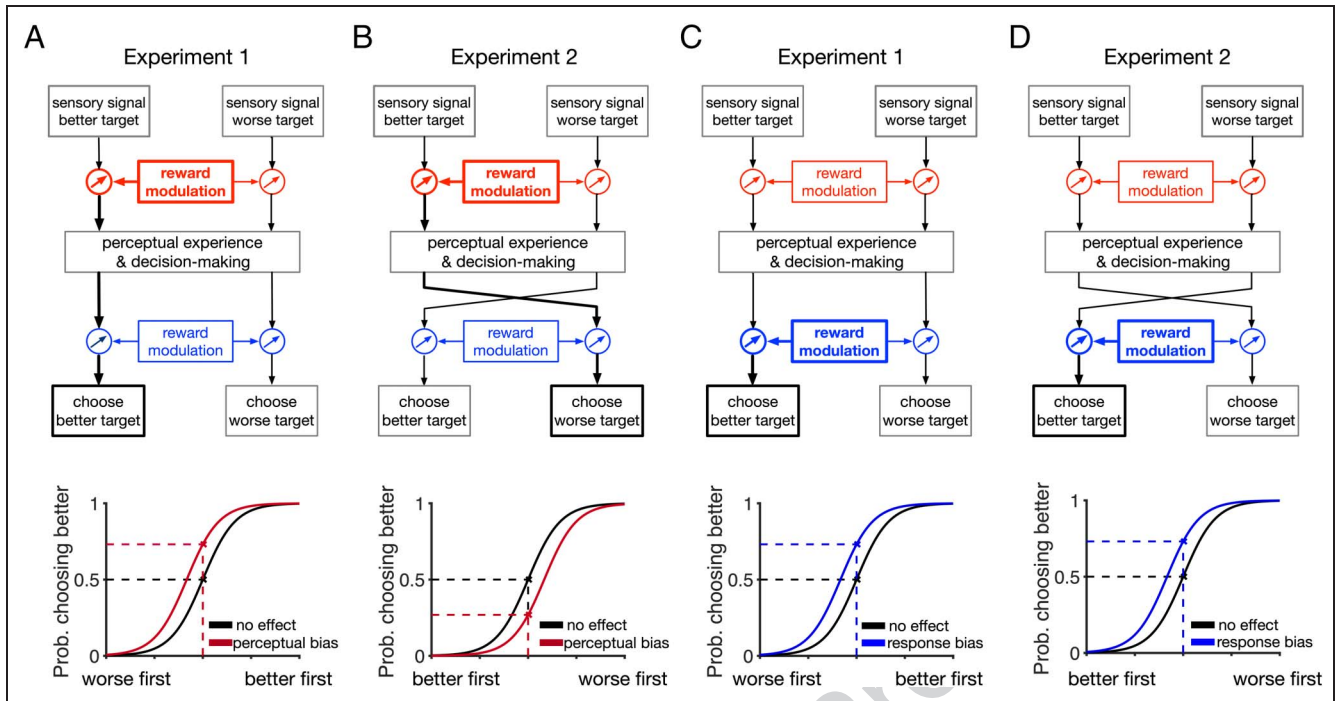


Figure 1. The effects of reward-induced changes in sensory processing and later stages of decision making on target selection in Experiments 1 and 2. (A–B) Modulation of sensory processing by reward information and its effect on target selection in Experiments 1 (A) and 2 (B). Reward information can enhance or facilitate processing of the target that appear on the more rewarding side (depicted with thicker lines) causing this target to be perceived earlier (perceptual bias). This would increase the probability of choosing the target on the more rewarding side in Experiment 1 (A) but increase the probability of choosing the target on the less rewarding side in Experiment 2, in which the participants have to saccade to the target that appeared second (B). The plots at the bottom depict changes in the probability of choosing the better target as a function of the TOA favoring the better (worse) target in Experiment 1 (respectively, Experiment 2). The black curve shows the probability of choosing the better target in the absence of any reward modulation. The highlighted arrow indicates the locus of reward modulation. (C–D) Modulation of later stages of decision making by reward information and its effect on target selection in Experiments 1 (C) and 2 (D). Reward modulation of later stages of decision making increases the probability of choosing the better target (response bias) in both experiments similarly. Conventions are the same as in A–B [R1.2].

participants, 21 performed in both Experiments 1 and 2, each of which consisted of four sessions. The remaining eight participants only performed in either Experiment 1 or 2. All participants gave informed consent to participate according to a protocol approved by the Dartmouth College Institutional Review Board. All participants signed a written consent form before participating in the experiments.

Experimental Design

To compare different mechanisms of how reward influences perceptual decision-making, we measured the effects of reward on target selection during a TOJ task. We used a within-participant design in which human participants performed two variations of a temporal judgment task known as the paired-target task under three different reward conditions (Neutral, Gain, and Loss; see Reward Conditions section). The participants were required to detect the order of two visually identical targets (gabor patches) that appeared on the computer screen at varying time intervals with respect to each other (Schiller & Chou, 1998). The participant's task was to saccade to the first target in Experiment 1 (after both

targets were presented), whereas in Experiment 2, the participant's task was to saccade to the second target (Figure 2). Participants were instructed to saccade to the correct target (first target in Experiment 1 and second target in Experiment 2) without being concerned about the speed (i.e., reward did not depend on response time). Following each response, participants received reward feedback—a green or red circle with the amount of reward points earned or lost, respectively (Figure 2).

To motivate the participants, participants' total reward points were exchanged for a monetary reward at the end of each experiment based on their performance. They were compensated with a combination of money and "t-points," extra credit points for classes within the Department of Psychological and Brain Sciences at Dartmouth College. More specifically, in addition to the base rate of \$10/hr or 1 t-point/hr, participants were compensated up to an additional \$10/hr depending on the total reward points they collected in each experiment. The order of two experiments was randomized across participants to avoid possible confounds.

To ensure consistent processing of the two targets, participants were required at the beginning of each trial to fixate on a white cross at the center of the screen for at

least 500 msec (see Figure 2 for the trial sequence). Fixation was considered broken when the participant's eye position deviated 112.5 pixels ($\sim 2^\circ$ visual angle) from the fixation cross before the fixation period ended. If the participant broke fixation, a new trial would begin after a 1000-msec pause. After successful fixation, the amounts of reward points to be earned/lost upon a correct/incorrect response to each target were signaled on the side of the fixation cross corresponding to that target.

The amounts of reward points expected from correct and incorrect responses, which we refer to as gains and losses, on both sides were presented close to the fixation cross and large enough to be read without breaking fixation ($\sim 2^\circ$ visual angle from the fixation cross). To present different reward information more distinguishably, the amounts of gains and losses were presented in green and red, respectively. After the offset of reward information at ~ 1000 msec, there was a variable interval between 500 and 1000 msec (uniform distribution) before the first target appeared on the screen. The second target then appeared after an interval selected from the following values: 0, 16.7, 33.3, 50, and 66.7 msec. We refer to these values as the TOA. Targets were presented at equal distances from the fixation cross ($\sim 7^\circ$ visual angle).

All stimuli were presented on an FSI AM250 monitor, which has a refresh rate of 60 Hz and resolution of 1920×1080 pixels. Participants were seated 60 cm from the computer screen. Eye movements were recorded using

a video-based eye-tracking system (Eyelink 1000, SR Research Ltd). To minimize head movements, participants were seated with their chin on a chin rest. The experiments were programmed using PsychToolbox in MATLAB (Kleiner et al., 2007; Brainard, 1997; Pelli, 1997).

Reward Conditions

Both Experiments 1 and 2 consisted of four sessions, each of which corresponded with one of the three reward conditions: Neutral, Gain, and Loss. In the Neutral condition, which always preceded either the Gain or Loss condition, correct and incorrect responses to either target resulted in gaining or losing 3 points, respectively (indicated by the [3, -3] on the two sides of the fixation point; Figure 2). In the Gain condition, reward points for saccade to one target were [5, -3], whereas the other target had reward points of [1, -3]. In the Loss condition, the targets were associated with reward points of [3, -1] and [3, -5]. These values were selected based on our pilot study to ensure similar differences in gain and loss values, assuming an average loss aversion factor of 2. During the Gain and Loss conditions, the side with the target with higher expected value was randomly assigned on each trial. The order of the Gain and Loss conditions was randomized across participants. Each reward condition consists of 180 trials and was performed in a single session of the experiment without pause (lasting

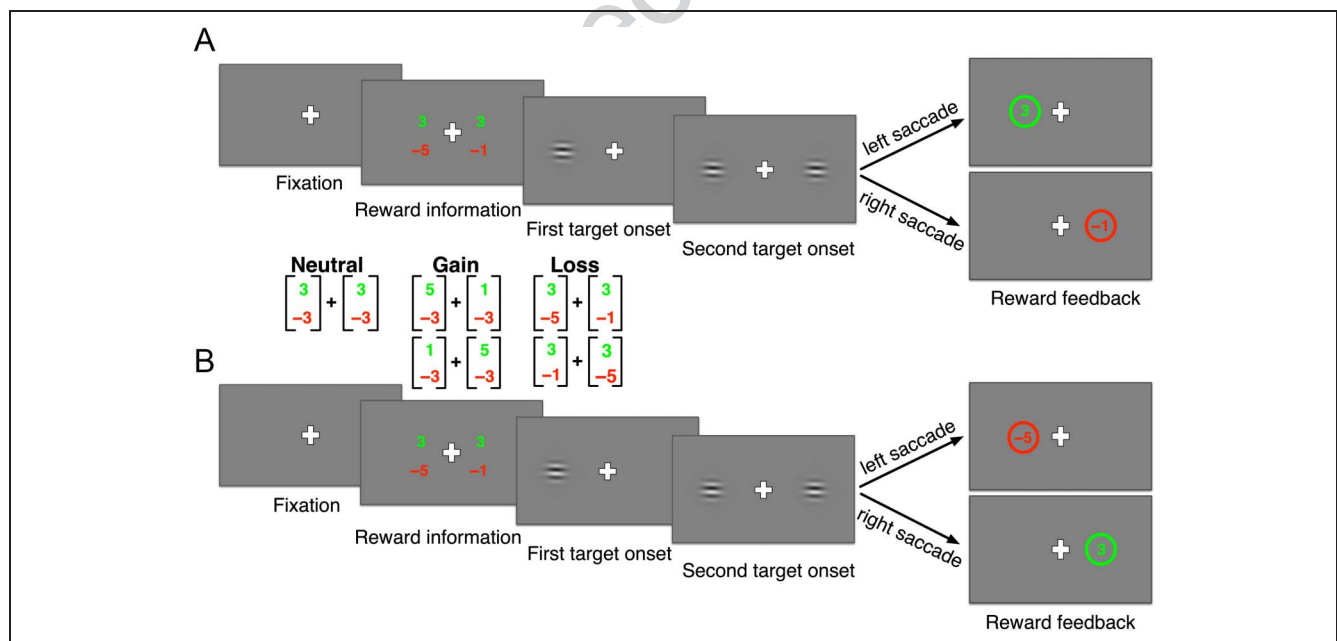


Figure 2. Schematic of the experimental paradigm. (A) Timeline of a trial during Experiment 1. Each trial began with a fixation cross followed by the presentation of reward points associated with a correct/incorrect response for each target on both sides of the fixation cross (in green for gains and red for losses). The amount of expected reward was manipulated in three experimental conditions (Neutral, Gain, and Loss) as indicated in the inset. Following the presentation of the reward information, two identical targets (gabor patches) appeared on the screen asynchronously. The participants' task was to report the first target that appeared on the screen by making a saccade to the target. Reward feedback was then given by a green circle for correct response or red circle for incorrect response around the selected choice with reward points gained or lost, respectively, in the center of the circle. (B) Timeline of a trial during Experiment 2. Stimulus and reward information were presented similar to Experiment 1 except the correct response required making a saccade to the second target that appeared on the screen.

about 15–20 min). Therefore, each session only involved one of the three reward conditions.

Fitting Choice Behavior and Estimated Parameters

We fit choice data from each session (reward condition) of the experiment separately to estimate different aspects of target selection. The psychometric function in Experiment 1 was defined as the probability of choosing the left target as a function of the TOA favoring the left target, $TOA = t_{\text{right}} - t_{\text{left}}$, where t_{left} and t_{right} are the onset time of the left and right targets, respectively. The psychometric function in Experiment 2 was defined as the probability of choosing the left target as a function of the TOA favoring the right target (by flipping the sign for the TOA on each trial) to measure target selection as a function of the signal relevant for the task. We then used the standard maximum likelihood estimation (by minimizing the negative log likelihood) to fit the psychometric function in each session using the following equation:

$$p_L(TOA) = \frac{1}{1 + e^{-(\beta_{\text{side}} + \beta_s \times TOA + c \times \beta_r)}} \quad (1)$$

where β_{side} is the side bias measuring the participant's preference for saccade to target on the left side of the fixation cross regardless of what was presented, β_s represents sensitivity to sensory information (i.e., how selection changes as a function of the TOA), β_r is the reward bias measuring the preference in target selection toward the side with the larger expected reward. Finally, c is a dummy variable that indicates the more rewarding side on each trial (i.e., it is 1 if the left side is more rewarding, 0 if both sides are equally rewarding, and -1 if the right side is more rewarding). Critically, these parameters capture different aspects of target selection and how this process is influenced by reward information. Note that the probability of choosing the left option in Equation 1 is dimensionless and thus β_{side} , $\beta_s \times TOA$, and β_r are also dimensionless. As a result, β_s has unit of $\frac{1}{\text{msec}}$ because TOA measures time in msec. Throughout the article, we report the dimensionless reward bias (β_r), side bias (β_{side}), and sensitivity that has the unit of msec^{-1} . We also report the TOA equivalent of shifts in target selection due to reward information and inherent bias in selection after dividing β_{side} and β_r by β_s [R2.m].

We also computed error in parameter estimation with two different methods. First, we used the Hessian matrix of the log-likelihood function to find the 95% confidence interval of the estimated parameters. Second, we randomly sampled 95% of the data 50 times to fit the ensuing psychometric function and then calculated the mean and standard deviation of the estimated parameters across all samples.

Data Exclusion

We excluded 20 sessions from the total 200 sessions of the two experiments completed by the 29 participants. Exclusion was performed on a session-by-session basis using three exclusion criteria. First, we excluded sessions in which sensitivity to the TOA was negative, indicating that the participant did not perform the task properly by ignoring the main task variable (TOA) on most trials. Second, we excluded sessions in which the overall task performance did not exceed chance (50%) plus two times *SEM*, reflecting an unusually poor performance. Using these two criteria, we removed 19 sessions from eight different participants (seven of these participants had one or two excluded sessions except one participant whose all eight sessions were excluded). Finally, we discarded sessions in which either of the fitting parameters (i.e., β_s , β_{side} , or β_r) deviated by more than three times the standard deviation from the corresponding parameter's mean across all sessions. The third criteria led us to remove one more session. Results reported here are based on the remaining 180 sessions (valid sessions).

Statistical and Data Analysis

We used the standard maximum likelihood estimation to find the parameters of the psychometric function in each session. Then, we used two-sided signed rank test to compare the estimated parameters with the null hypothesis (0 corresponding to no effect). To compare reward biases between the Gain and Loss conditions, we used two-sided Wilcoxon rank-sum test. We used both Pearson and Spearman correlation to examine the correlation between estimated parameters.

To determine the model that best explains the variances in the saccadic RT (SRT), we used a stepwise general regression model (GLM). We included the following regressors in the stepwise GLM: unequal reward condition indicating unequal (Gain and Loss) or equal (Neutral) expected reward outcomes, the TOA, response accuracy (correct vs. incorrect response), and a dummy variable indicating whether the chosen target had the higher or lower expected reward (chosen-target relative value). The last regressor only applied to the Gain and Loss conditions. A stepwise GLM procedure examines all combinations of regressors and their interactions to determine terms whose inclusion results in a significant increase of the adjusted R^2 . We used custom codes and the statistical package in MATLAB (MathWorks, Inc.) to perform all simulations and statistical analyses.

Optimality Analysis

The optimal reward bias is defined as the amount of shift in target selection due to reward information that maximizes the total reward earned in a given session. To determine the optimal reward bias, we first calculated the

expected amount of reward earned assuming a given level of sensitivity to sensory information and loss aversion. The participant's sensitivity can predict the overall number of correct choices, and thus, the number of reward points they can earn. Loss aversion causes individuals to respond more strongly to reward points lost than gained (note that reward points were assigned to the two sides on a trial-by-trial basis). Considering these factors, the expected payoff associated with saccades to the left (L) and right (R) can be calculated using the following equations:

$$\begin{aligned} \text{payoff}_L(p, \gamma) &= (p(\beta_s, \text{TOA}) \times \text{gain}_L) \\ &\quad + (\gamma(1 - p(\beta_s, \text{TOA})) \times \text{loss}_L) \\ \text{payoff}_R(p, \gamma) &= ((1 - p(\beta_s, \text{TOA})) \times \text{gain}_R) \\ &\quad + (\gamma p(\beta_s, \text{TOA}) \times \text{loss}_R) \end{aligned} \quad (2)$$

where γ represents the loss aversion factor and $p(\beta_s, \text{TOA})$ is the probability of the correct response for a given TOA:

$$p(\beta_s, \text{TOA}) = \frac{1}{1 + e^{-(\beta_s \times \text{TOA})}} \quad (3)$$

where β_s represents sensitivity to sensory information. The total expected payoff for a given value of shift in target selection, μ_r , is equal to

$$\begin{aligned} \text{payoff}(\text{TOA}, \mu_r, \beta_s, \gamma) &= P_{\text{cho}}(\text{TOA}, \mu_r, \beta_s) \\ &\quad \times \text{payoff}_L \\ &\quad + (1 - P_{\text{cho}}(\text{TOA}, \mu_r, \beta_s)) \\ &\quad \times \text{payoff}_R \end{aligned} \quad (4)$$

where $P_{\text{cho}}(\text{TOA}, \mu_r, \beta_s)$ represents the probability of choice for a given TOA, μ_r , and β_s , which is computed as follows:

$$P_{\text{cho}}(\text{TOA}, \mu_r, \beta_s) = \frac{1}{1 + e^{-(\beta_s \times \text{TOA} + \mu_r)}} \quad (5)$$

As a result, the total expected amount of reward earned for all values of the TOA for given values of μ_r and β_s is equal to

$$ER(\mu_r, \beta_s, \gamma) = \int \text{payoff}(\text{TOA}, \mu_r, \beta_s, \gamma) d\text{TOA} \quad (6)$$

where the integral is computed by summing over all values of the TOA. Finally, the optimal reward bias can be computed by finding a value of μ_r that maximizes the value of ER for specific values of sensitivity and loss aversion factor (using Equation 6). We compared reward biases of individual participants with the predicted

optimal values ($=\mu_{\text{r opt}}$) to examine the deviations from optimality.

Computational Model

The basic model consisted of two cortical columns with two pools of excitatory neurons and one inhibitory pool of neurons in the superficial layer and two pools of excitatory neurons in the deep layer (Soltani, Noudoost, & Moore, 2013; Supplementary Figure 1). All neural pools in all the layers received a background input mimicking input from adjacent cortical neurons with different types of selectivity. The excitatory pools in the superficial layer also received visual input related to the presentation of targets on the screen. Specifically, the visual input to the two pools were similar except that they had different onset timing according to the TOA on each trial. Moreover, the two pools of excitatory neurons in the superficial layer were mutually inhibited using a shared pool of inhibitory interneurons. This mutual inhibition created a winner-take-all competition and caused activity in excitatory pools to diverge in both the superficial and deep layers. We used a mean field approximation of a spiking network model of decision-making to simulate the superficial layer (Wong & Wang, 2006). Each excitatory pool of neurons in the deep layer had weak self-excitatory recurrent connections. The deep layer then projected its output to the brain stem or superior colliculus to direct a saccadic eye movement. We determined the choice of the network on each trial by identifying the first deep-layer excitatory pool whose activity passed 15 Hz (considered the winner pool). The full details of the basic model are described elsewhere (Soltani et al., 2013).

To simulate the observed effects of unequal reward information, we considered three alternative mechanisms. These mechanisms affected different parts of the model, mimicking either modulations of sensory processing or later stages of decision-making processes. First, to simulate the effect of reward information on later stages of decision-making processes, we included a reward-based input to the excitatory pools in the deep layer (Figure 9A; Mechanism 1). This input was independent of the amount and timing of the visual input (TOA) on a given trial. Second, to simulate the effect of reward information on sensory processing, we assumed two alternative mechanisms in which reward information could modulate the visual input to the decision circuit. In the first mechanism, the input evoked by targets with a larger and smaller expected reward was multiplied by $(1 + \lambda)$ and $(1 - \lambda)$, respectively, where λ is a constant that measures the modulation of visual input by reward information (Figure 9B; Mechanism 2). This results in stronger input for the target on the more rewarding side. In the second mechanism for modulating sensory processing, the input for the target with larger expected reward was modulated through a shift in TOA in favor of the target with larger expected reward (Figure 9B; Mechanism 3). This shift mimics faster processing of input related to the target

on the more rewarding side in higher visual areas. In both Mechanisms 2 and 3, the reward cue presented to participants modulates the processing of sensory information and thus could result in perceptual bias. Finally, to simulate Experiment 2 (selecting the second target that appears on the screen), we assumed that the projections of the output in excitatory pools of the superficial layer is switched via a gating mechanism (green dashed lines in Figure 9A, B) to allow the selection of the nonwinner pool as the response.

Because of the nonlinear dynamics of the proposed network models, we had no a priori predictions about which alternative mechanism would be more compatible with the observed lack of correlation between shifts in target selection and the sensitivity to the TOA, and whether a single mechanism was sufficient to capture our main experimental findings. Also, the two proposed mechanisms of reward influence on sensory processing—Mechanism 2 corresponding to enhancement and Mechanism 3 to facilitation of sensory signals due to unequal reward—were different enough to warrant the simulation and examination of both mechanisms.

RESULTS

Alternative Mechanisms for the Influence of Reward on Target Selection and Their Predictions

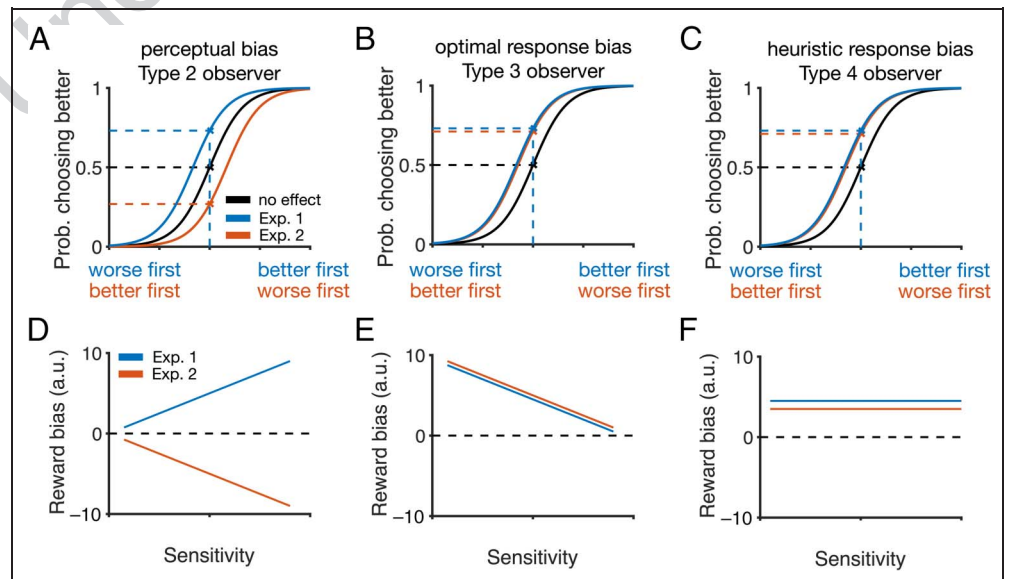
To study the influence of reward information on perceptual decision-making in general and in temporal judgment in particular, we used modified versions of the paired-target task in which the participants reported what they perceived to be the first (Experiment 1) and second (Experiment 2) of the two targets to appear on the screen

by making a saccade to the target. At the beginning of each trial, the amounts of reward points expected to be gained and lost upon correct and incorrect responses, respectively, were presented on the two sides of the fixation cross and were manipulated across experimental sessions (Figure 2; see Methods).

As described earlier, because reward information is not informative about the correct response, an observer could ignore reward information or use it to enhance sensory processing on both sides equally (i.e., similar to the effect of arousal). In either case, such observers (Type 1 observers) would not demonstrate any shift in choice toward the better or worse target. Shifts in target selection due to reward (i.e., reward bias), however, could happen through different but nonexclusive mechanisms. First, reward information could bias attention to the better side, resulting in enhanced processing of visual information on that side and thus perceptual bias in Type 2 observers (Figure 1A, B). Second, reward information could bias later stages of decision-making processes toward selection of the better target, causing response bias (Figure 1C, D) that could be dependent or independent of sensitivity to sensory information (Types 3 and 4 observers). Importantly, these different mechanisms for the influence of reward result in different patterns of reward bias in the two experiments: Changes in sensory processing would result in opposite shifts in target selection in Experiments 1 and 2, but changes in later stages of decision-making would similarly shift target selection in the two experiments (Figure 3A–C) [R1.2].

In addition to the pattern of the shifts in target selection, the relationship between these shifts and sensitivity to sensory information may be used to further distinguish between alternative mechanisms. First, the reward-induced

Figure 3. Reward bias and its relationship to sensitivity to sensory information as predicted by different mechanisms for the influence of reward. (A) In Type 2 observer, bias in target selection caused by differential processing of the two targets results in opposite shifts in target selection in the two experiments. Plot depicts a hypothetical psychometric function for a Type 2 observer that exhibits perceptual bias. The blue and orange labels for the x-axis correspond to the Experiments 1 and 2, respectively. Conventions are the same as in Figure 1. (B) A Type 3 observer with response bias exhibit similar shifts in target selection in the two experiments. (C) A Type 4 observer with response bias exhibit similar shifts in target selection in the two experiments. (D–F) Predictions of correlation between reward bias and sensitivity to sensory information for a Type 2 (D), Type 3 (E), and Type 4 (F) observers.



attentional effects on sensory processing could cause larger shifts in target selection for decision makers who are more sensitive to sensory signal (i.e., participants with larger sensitivity). This would predict a specific pattern for reward bias as a function of sensitivity for Type 2 observers (Figure 3D) [R1.6]. In contrast, decision makers that optimize their shift in target selection due to reward information (Type 3 observers) would show a decrease in reward bias as a function of sensitivity to sensory information (Figure 3E) [R1.4].

To illustrate this point, we computed the optimal reward bias based on given values of sensitivity to sensory evidence (β_s) and loss aversion factor (γ) separately in the Gain and Loss conditions (Figure 4A, D; see Optimality Analysis section in Methods). We found that, in the Gain condition, the optimal bias should decrease with larger sensitivity (Figure 4B). Moreover, the optimal reward bias should decrease as loss aversion increases for a given level of sensitivity (Figure 4C). In the Loss condition, the optimal reward bias should also decrease with larger sensitivity (Figure 4E) but increase with larger loss aversion (Figure 4F). Therefore, these results show that optimal shift in target selection requires reward bias to be inversely correlated with the individual's level of sensitivity to sensory evidence. Moreover, loss aversion should have opposite effects on reward bias in the Gain and Loss conditions.

Together, these results predict that the effect of reward on early sensory processing and perceptual bias (Type 2 observers) can be detected from opposite shifts in target selection in Experiments 1 and 2 (Figure 3A) whereas the effect of reward on later stages of decision-making and ensuing response bias (Types 3 and 4 observers) could be

revealed from similar shifts in target selection in the two experiments (Figure 3B–C). In addition, shifts in target selection in observers that optimize response bias (Type 3 observers) depend on sensitivity to sensory information; shifts should be small when sensitivity is high (corresponding to good temporal judgment) and large if sensitivity is low, corresponding to poor temporal judgment (Figure 3E). On the other hand, in Type 4 observers, the shifts in target selection are independent of sensory information (Figure 3F) [R1.4].

Reward Information Affects Target Selection

We used our experimental data to test alternative predictions about the effects of reward on target selection depicted in Figure 3 to identify mechanisms by which reward influences TOJ. To that end, we fit each participant's psychometric function (the probability of choosing the left target as a function of the TOA) using a sigmoid function to estimate three parameters for each participant: sensitivity to sensory information, side bias, and reward bias (see Methods for more details). Sensitivity measures the fidelity of target selection to sensory evidence (i.e., TOA), side bias measures an overall bias in choosing the left or right target independently of sensory information, and reward bias measures bias in target selection toward the more rewarding side (Figure 5A–C). These parameters were estimated separately for each of the three reward conditions.

We found that participants were sensitive to the TOA in both experiments. The average values ($\pm SD$) of sensitivity in Experiment 1 were equal to 0.0432 ± 0.0191 , 0.0472 ± 0.0222 , and 0.0438 ± 0.0179 (msec^{-1}) for the Neutral,

Figure 4. Predicted optimal shift in target selection due to unequal expected reward. (A) The expected payoff (reward points) for a given level of sensitivity ($\beta_s = 0.05 \text{ msec}^{-1}$) and reward bias during the Gain condition. Each curve represents the expected payoff for a different level of loss aversion (loss aversion factor γ). The peak in each curve indicates the optimal reward bias. (B) The optimal reward bias as a function of the sensitivity to sensory evidence. The optimal reward bias diminishes as sensitivity increases. (C) The optimal reward bias as a function of the loss aversion factor in the Gain condition separately for three different values of the sensitivity. The optimal reward bias decreases as loss aversion increases in the Gain condition. (D–F) The same as in A–C but for the Loss condition. Similar to the Gain condition, the optimal reward bias decreases with larger values of sensitivity. However, the optimal reward bias increases as loss aversion increases in the Loss condition.

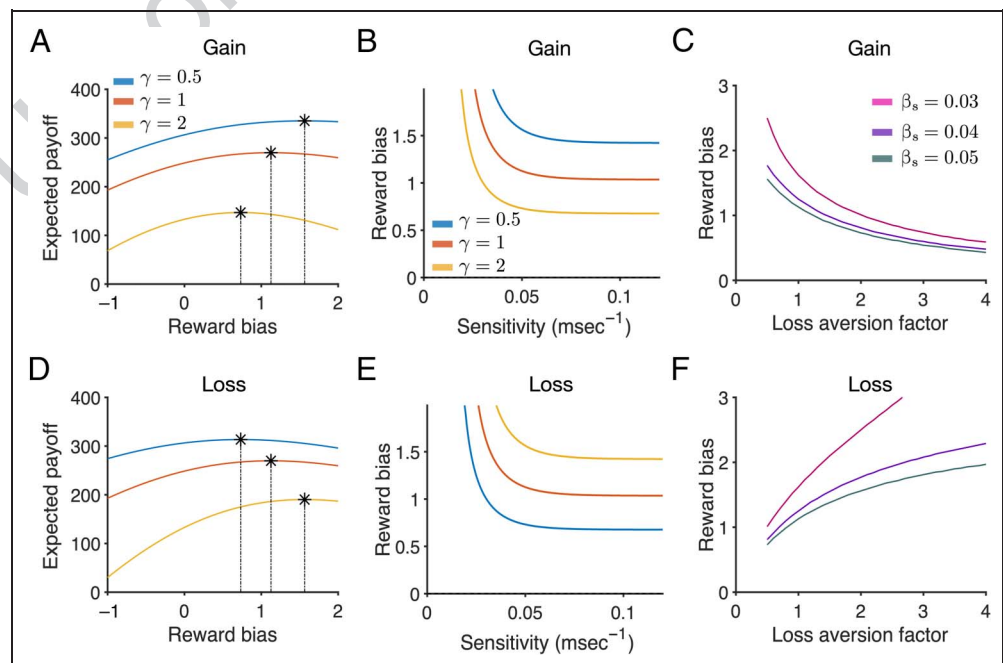
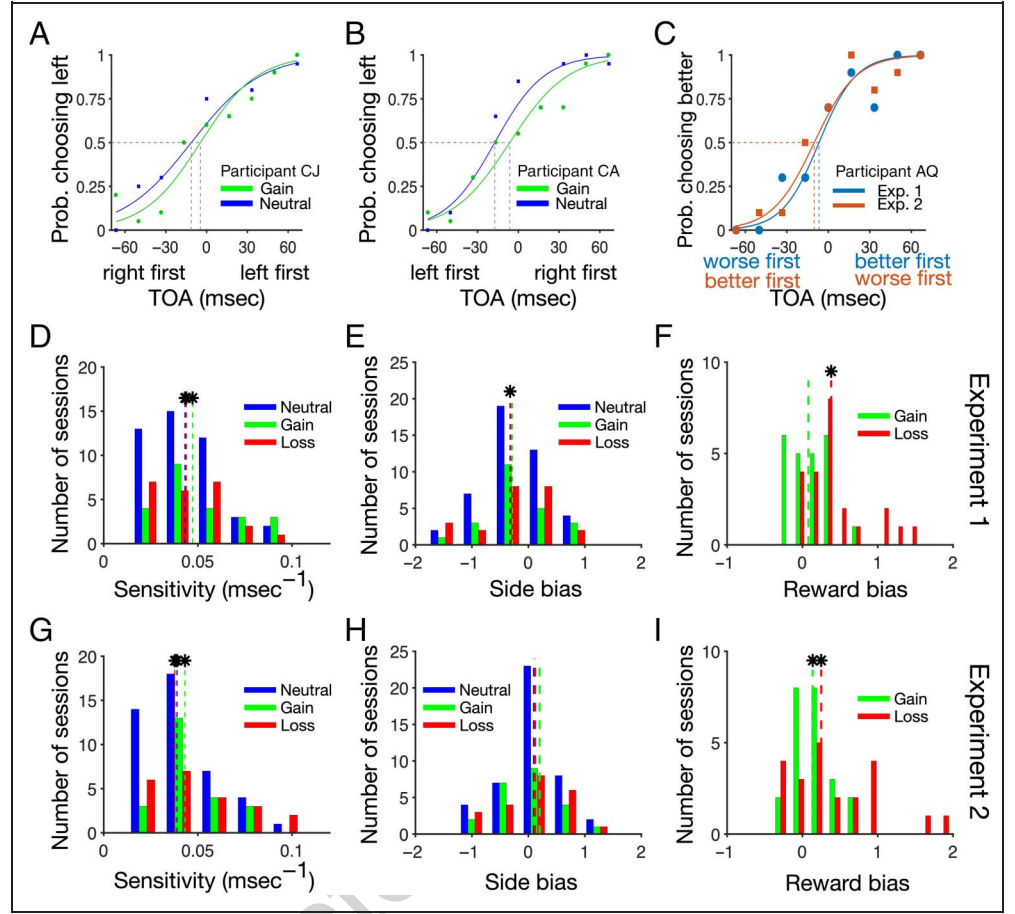


Figure 5. Influences of unequal expected reward on target selection. (A–C) Example psychometric functions in different conditions of Experiments 1 and 2. (A) Plotted is the probability of choosing the left target as a function of the TOA for an example participant in the Gain and Neutral conditions of Experiment 1. This participant exhibited a slightly larger level of sensitivity in the Gain compared with Neutral condition. Positive (negative) values of TOA correspond to the left (right) target to appear first. Dashed lines indicate the TOA at which the two targets are selected equally. (B) The same as in A but for another participant who exhibited a larger side bias in the Neutral than in Gain condition of Experiment 2. In Experiment 2, negative (positive) values of TOA correspond to the left (right) target appearing first. (C) Psychometric functions of an example participant in the Gain condition of Experiments 1 and 2. Plotted is the probability of choosing the better target as a function of the TOA for that target. For Experiment 1, negative (positive) values of TOA means the target with smaller (larger) expected reward has appeared first. In Experiment 2, negative (positive) values of TOA means the target with smaller (larger) expected reward appeared second. This participant exhibited similar reward bias in the two experiments. (D–F) Histograms plot the number of participants with given values of sensitivity (D), side bias (E), and reward bias (F) under different reward conditions during Experiment 1. The dashed lines show the medians, and each asterisk indicates a significant difference from 0 (two-sided signed-rank test, $p < .05$). (G–I) The same as in D–F but for Experiment 2.



Gain, and Loss conditions, respectively, and significantly larger than zero (two-sided signed-rank test; Neutral: $p = 5.18 \times 10^{-9}$, $d = 2.25$; Gain: $p = 2.7 \times 10^{-5}$, $d = 2.29$; Loss: $p = 2.7 \times 10^{-5}$, $d = 2.47$; Figure 5D). The average values of sensitivity in Experiment 2 were equal to 0.0389 ± 0.0195 , 0.0432 ± 0.0187 , and 0.0378 ± 0.0259 (msec⁻¹) for the Neutral, Gain, and Loss conditions, respectively. As in Experiment 1, these average values were significantly larger than zero (two-sided signed-rank test; Neutral: $p = 7.6 \times 10^{-9}$, $d = 2.08$; Gain: $p = 2.7 \times 10^{-5}$, $d = 2.34$; Loss: $p = 4 \times 10^{-5}$, $d = 1.76$; Figure 5G).

We compared sensitivity across the three conditions in Experiments 1 and 2 to test possible differences between the two experiments in terms of the overall task difficulty. However, we found no significant difference in sensitivity over all conditions of the two experiments (two-sided signed-rank test; $p = .26$, $d = 0.06$). In addition, there was no significant difference in sensitivity in either of the three conditions between the two experiments (two-sided signed-rank test; Neutral: $p = .44$, $d = 0.06$; Gain: $p = .36$, $d = 0.17$; Loss: $p = .81$, $d = 0.03$). Thus, we did not find any evidence for a difference in task difficulty (as

measured by sensitivity) between the two experiments [R2.1].

We also examined estimated side bias in different reward conditions and experiments. We note that, to minimize possible side bias, the side associated with the target with better reward outcomes had been randomly assigned for each trial. In Experiment 1, we did not observe any evidence for side bias except in the Neutral condition (two-sided signed-rank test; Neutral: -0.32 ± 0.52 , $p = .003$, $d = 0.48$; Gain: -0.28 ± 0.53 , $p = .094$, $d = 0.39$; Loss: -0.31 ± 0.70 , $p = .114$, $d = 0.36$; Figure 5E). Furthermore, we did not find any evidence for side bias in any conditions of Experiment 2 (two-sided signed-rank test; Neutral: 0.10 ± 0.49 , $p = .37$, $d = 0.06$; Gain: 0.20 ± 0.57 , $p = .93$, $d = 0.03$; Loss: 0.12 ± 0.64 , $p = .91$, $d = 0.03$; Figure 5H). We also found similar results when we measured side bias in terms of the TOA (Supplementary Figure 2A, C). Together, these results illustrate that participants exhibited very small side bias in both experiments.

Having established that participants performed the temporal judgment task appropriately without significant

side bias, we then examined reward bias measuring the effect unequal expected reward on target selection. In Experiment 1, 87% and 96% of participants exhibited a significant reward bias in the Gain and Loss conditions, respectively. However, across all participants, reward bias toward the more rewarding side was significant only in the Loss condition (two-sided signed-rank test; Gain: 0.08 ± 0.24 , $p = .107$, $d = 0.38$; Loss: 0.38 ± 0.44 , $p = 1.44 \times 10^{-4}$, $d = 0.99$; Figure 5F). In Experiment 2, 83% and 96% of participants exhibited a significant reward bias in the Gain and Loss conditions, respectively, and across all participants, reward biases were significantly larger than zero in both conditions (two-sided signed-rank test; Gain: 0.13 ± 0.24 , $p = .024$, $d = 0.55$; Loss: 0.25 ± 0.61 , $p = .008$, $d = 0.65$; Figure 5I). We found similar results when we measured reward bias in terms of the TOA (Supplementary Figure 2B, D). Together, our results show that most participants used reward information to bias their target selection. Hence, our findings are not compatible with Type 1 observers.

To ascertain that noise in the estimation of the parameters did not influence our results, we calculated the error of estimation based on two different methods (see Methods for more details). Using the first method, the Hessian matrix of the log-likelihood function, we found small errors (mean $\sim 14\%$) in the estimation of reward biases (Supplementary Figure 3A). The second method that was based on resampling produced a slightly larger error estimation (mean $\sim 20\%$; Supplementary Figure 3B). These results demonstrate the robustness of our fitting procedure.

Reward Effects Are Mediated through Changes in Later Stages of Decision-making

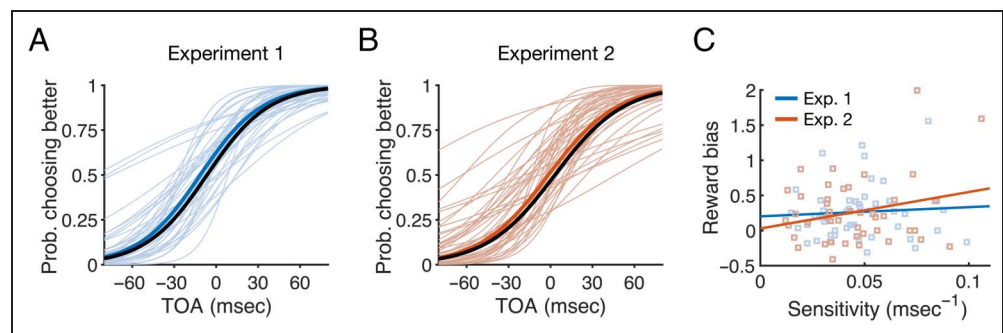
To test the predictions of alternative mechanisms, we next compared reward bias in the two experiments. As detailed above, a difference between shifts in target selection (or reward bias) in Experiments 1 and 2 would indicate reward effects on early stages of sensory processing and perceptual bias. However, we did not find a significant difference in reward biases between Experiments 1 and 2 on

average (Wilcoxon rank-sum test; $p = .69$, $d = 0.004$; Figure 6A, B) or for the Gain or Loss condition separately (Wilcoxon rank-sum test; Gain: $p = .645$, $d = 0.08$; Loss: $p = .594$, $d = 0.04$; Figure 5F, I). This was also true when comparing reward biases in individuals who performed both experiments successfully (two-sided signed-rank test; Gain: $\Delta = -0.04 \pm 0.39$, $p = .44$, $d = 0.15$; Loss: $\Delta = 0.07 \pm 0.71$, $p = .76$, $d = 0.03$) [R2.1]. Therefore, our results do not provide evidence for reward effects on early stages of sensory processing and perceptual bias and instead are more compatible with response bias.

Although both Type 3 and Type 4 observers exhibit similar shifts in Experiments 1 and 2, reward biases in Type 3 but not Type 4 observers were correlated with sensitivity to sensory information (Figure 3E, F). Therefore, we examined the correlation between reward bias and sensitivity of individual participants in the two experiments. However, we did not find any evidence for the correlation neither between reward bias and an individual's sensitivity in any conditions of the two experiments (Pearson correlation; Gain condition: Experiment 1: $r = .02$, $p = .91$; Experiment 2: $r = -.09$, $p = .66$; Loss condition: Experiment 1: $r = .27$, $p = .20$; Experiment 2: $r = .36$, $p = .10$) nor across the two conditions of each experiment (Pearson correlation; Experiment 1: $r = .07$, $p = .66$; Experiment 2: $r = .24$, $p = .10$; Figure 6C [R1.3]). This result suggests that the observed shift in target selection due to the reward information is more compatible with Type 4 observers.

Finally, we also compared reward bias between the Gain and Loss conditions. Although we did not find any evidence for correlation between reward bias in Gain and Loss conditions for participants who successfully performed (valid sessions) in both conditions (Pearson correlation; Experiment 1: $r = -.31$, $p = .22$, $n = 20$; Experiment 2: $r = -.02$, $p = .94$, $n = 17$; Spearman correlation; Experiment 1: $r = -.04$, $p = .88$, $n = 20$; Experiment 2: $r = .07$, $p = .79$, $n = 17$), there was an overall larger bias in the Loss than Gain condition in the two experiments (two-sided signed-rank test; $\Delta = 0.34 \pm 0.60$, $p = .006$, $d = 0.52$, $n = 37$; Supplementary Figure 4). The larger effect of unequal loss on choice behavior in our temporal judgment

Figure 6. Similar shifts in target selection in Experiments 1 and 2 and lack of correlation between reward bias and sensitivity across individual participants. (A) Plotted are the psychometric functions of individual participants (light blue) and median psychometric function (solid blue curve) across all participants in Experiment 1. The solid black curve shows the psychometric function in absence of any reward biases. Data from both the Gain and Loss conditions are included. (B) Similar to A, but for Experiment 2. (C) Plotted is the reward bias versus sensitivity across individual participants during Experiments 1 and 2. The solid blue and orange lines show the linear fit to the reward bias and sensitivity data points in Experiments 1 and 2, respectively [R1.3].



task resembles the well-known phenomenon that losses have stronger impact on choice behavior than gains of similar size, thus providing evidence for loss aversion in perceptual decision-making with possible gains and losses [R2.2].

Comparison of Participants' Reward Bias with Optimal Values

As we showed above, our experimental results are more compatible with response bias that was determined heuristically (Type 4 observer) and not based on optimization (Type 3 observer). To make this point more directly, we also compared the observed and optimal values of reward bias for individual participants assuming different values of loss aversion (Figure 4). We found that in the Gain condition of both experiments, participants exhibited reward biases that were smaller than the predicted optimal biases based on loss neutrality, corresponding to loss aversion factor equal to 1 (two-sided signed-rank test; $\Delta = -1.19 \pm 0.06$, $p = 3.5 \times 10^{-9}$, $d = 3.06$; Figure 7A–C). The observed reward biases would be optimal only if loss aversion factor was very large because larger loss aversion gives rise to smaller shifts in the Gain condition.

We also found the observed reward biases in the Loss condition to be smaller than the optimal values based on loss neutrality (two-sided signed-rank test; $\Delta = -1.11 \pm 0.10$, $p = 6.6 \times 10^{-8}$, $d = 1.39$; Figure 7D–F). In this case, however, the observed reward biases would be optimal if loss aversion factor was very small because smaller loss aversion gives rise to smaller shifts in Loss condition. Therefore, the observed smaller-than-optimal

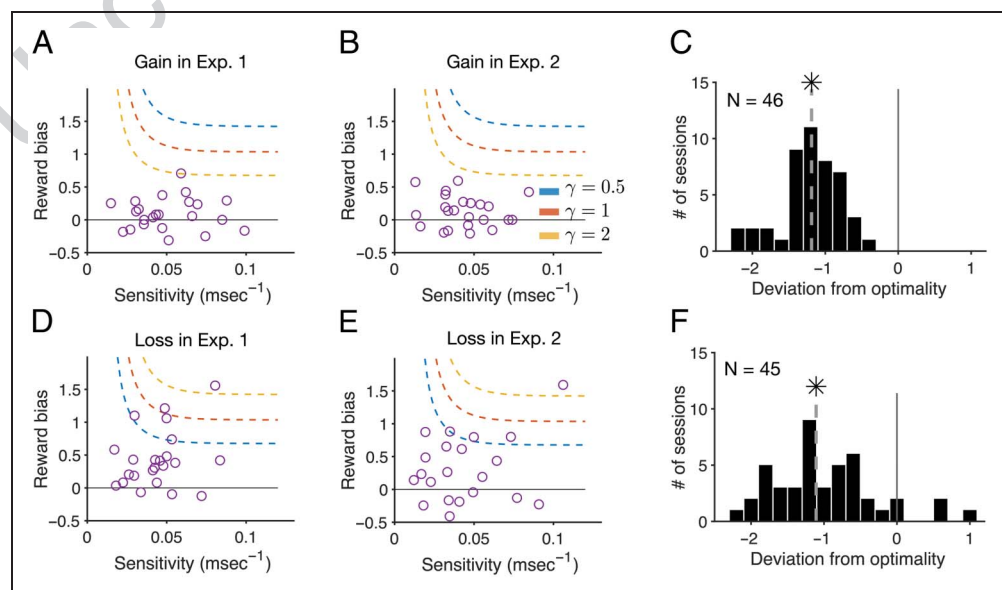
shifts in the Loss condition point to strong loss-seeking as opposed to strong loss-averse behavior that is seen in the Gain condition. Together, these results illustrate that the amount of shift in target selection due to unequal expected reward was suboptimal. As demonstrated below, our modeling results can explain why such optimization is not possible because of the loci of reward influence.

RT Reflects the Effect of Task Parameters

In our experiments, the participants were not instructed to saccade as quickly as possible and had to wait until both targets were presented before making a saccade. Nonetheless, we analyzed the SRT using a stepwise GLM model (see Methods) to examine whether the SRT reflects any task parameters. The stepwise GLM revealed that the TOA, unequal reward condition, response accuracy, and interaction between the TOA and response accuracy and between the TOA and unequal reward condition had significant effects on the SRT (stepwise GLM: $F(5, 32394) = 452$, $p = 10^{-273}$, adjusted $R^2 = .065$).

First, we found that the SRT decreased with the absolute value of the TOA corresponding to easier trials (β for TOA = -1.11 , $p = .04$; Figure 8A). Second, unequal reward outcomes resulted in an overall decrease in the SRT in the Gain and Loss conditions compared with the Neutral condition (β for reward condition = -0.11 , $p = 1.03 \times 10^{-26}$; Figure 8A). Third, the SRT was significantly smaller for correct trials compared with incorrect trials (β for response accuracy = -0.076 , $p = .0005$; Figure 8B). As mentioned above, the stepwise GLM did not reveal a significant effect of the chosen-target relative value on the

Figure 7. The amount of shift in target selection was suboptimal in the Gain and Loss conditions of both experiments. (A) Plot shows individuals' reward bias as a function of their sensitivity to the TOA. The blue, red, and yellow lines represent the optimal value of reward bias for different values of loss aversion factor. (B) The same as in A but for Experiment 2. (C) Plotted is the distribution of the differences between observed and predicted optimal reward biases based on loss neutrality ($\gamma = 1$) across all participants in the Gain condition of Experiments 1 and 2. (D–F) The same as in A–C but for the Loss condition in the two experiments. Note that a larger loss aversion factor predicts larger reward bias in the Loss condition (the opposite is true for the Gain condition).



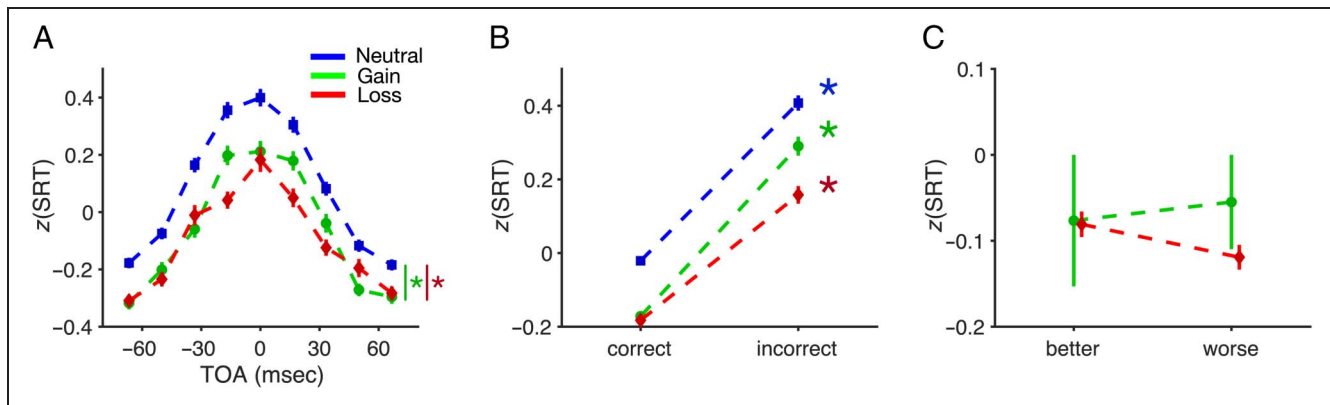


Figure 8. SRT was sensitive to both the TOA and unequal expected reward outcome, reflecting response accuracy but did not differ between the selection of the better and worse targets. (A) Plotted is the z-scored SRT as a function of the TOA for the Neutral (blue), Gain (green), and Loss (red) conditions. An asterisk shows a significant difference between the average SRT in the Neutral and Gain, or Neutral and Loss conditions (stepwise GLM, $p < .05$). (B) Plotted is the average z-scored SRT on correct and incorrect trials, separately for the three reward conditions indicated in A. An asterisk shows a significant difference between the average SRT on correct and incorrect trials (stepwise GLM, $p < .05$). (C) Plotted is the average SRT on trials in which the chosen target was the target with higher or lower expected reward corresponding to the better and worse targets, respectively.

SRT (Figure 8C). This lack of evidence for a significant effect could be caused by a few factors: (1) the stronger effects of TOA and response accuracy on the SRT, (2) different heuristics used to process gain and loss information, and (3) there was no time pressure in our experiments. Overall, these results show that SRT was sensitive to the TOA and unequal expected reward outcomes, indicating that both types of information influenced perceptual choice.

Plausible Neural Mechanisms for Observed Shifts in Target Selection

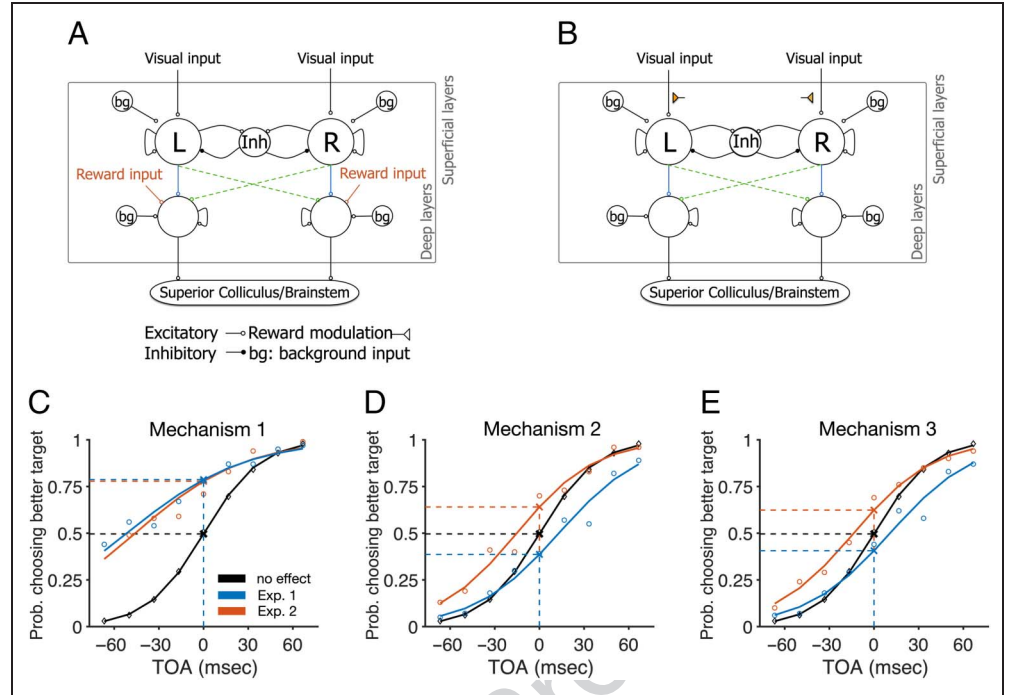
To reveal plausible neural mechanisms underlying the shifts in target selection, we simulated our experimental observations using a cortical network model that we have previously used to successfully simulate the paired-target task (Soltani et al., 2013). Specifically, we focused on capturing our two main experimental findings: (1) similar shifts in target selection during Experiments 1 and 2 and (2) lack of correlation between reward bias and individuals' sensitivity to sensory evidence in both experiments.

The model consisted of two neural columns with two pools of excitatory neurons and one inhibitory pool of neurons in the superficial layer, and two pools of excitatory neurons in the deep layer (Supplementary Figure 1; see Methods for more details). To simulate the effect of reward information, we considered three alternative mechanisms that could influence different parts of the model to mimic different stages of decision-making processes: reward-based input to the excitatory pools in the deep layer (Figure 9A; Mechanism 1); reward-dependent gain modulation of sensory input that gives rise to stronger input for the target on the more rewarding side (Figure 9B; Mechanism 2); and facilitation of response to the target with higher expected reward in higher visual areas (Figure 9B; Mechanism 3).

Simulation results showed that unequal expected reward results in significant reward bias using all three mechanisms (Figure 9C–E). These reward biases, however, were similar for Experiments 1 and 2 only in the model based on Mechanism 1 (Figure 9C). We also examined the correlation between reward bias and the sensitivity to visual input for a given set of model parameters. To generate target selection with different levels of sensitivity to sensory evidence, we changed the background noise in the input to the superficial layer. We did not find any evidence for correlation between reward bias and sensitivity to the TOA in the model based on Mechanism 1 (Pearson correlation; Experiment 1: $r = .07$, $p = .47$; Experiment 2: $r = .16$, $p = .12$; Figure 10A, D). This non-significant correlation, however, had a positive sign similar to that of the experimental data. In contrast, in the model based on Mechanism 2, reward bias were positively and negatively correlated with sensitivity to the TOA in Experiments 1 and 2, respectively (Pearson correlation; Experiment 1: $r = .89$, $p = 1.72 \times 10^{-35}$; Experiment 2: $r = -.89$, $p = 2.60 \times 10^{-35}$; Figure 10B, E). The model based on Mechanism 3 showed similar behavior to that of the model based on Mechanism 2 (Pearson correlation; Experiment 1: $r = .90$, $p = 3.73 \times 10^{-35}$; Experiment 2: $r = -.88$, $p = 7.03 \times 10^{-34}$; Figure 10C, F).

Overall, these results illustrate that the model based on Mechanism 1 is more compatible with our experimental data for two reasons: (1) exhibiting equal shifts in target selection in Experiments 1 and 2, replicating the observed response bias, and (2) lack of correlation between reward bias and sensitivity to the TOA in both experiments. These modeling results support the conclusion that the observed shift in participants' behavior due to the reward information is more likely to be due to changes in later stages of decision-making. In addition, the modeling results provide a plausible mechanism for how reward information influences perceptual choice. Finally, by

Figure 9. Alternative mechanisms for simulating the effects of unequal expected reward on perceptual decision making. (A) The extended model with independent reward input (Mechanism 1). In this model, excitatory pools in deep layer receive an additional reward-based input that was independent of the visual input. The blue solid lines for Experiment 1 and green dashed lines for Experiment 2 show the output projection of excitatory pools of superficial layer to deep layer excitatory pools. (B) The extended model with reward-dependent modulation of visual input (Mechanisms 2 and 3). In this model, the visual input to the excitatory pools of superficial layer is modulated by reward information provided at the beginning of the trial. This modulation was performed via two different mechanisms. In a



model with Mechanism 2, reward information strengthens (weakens) the visual input for the target with higher (lower) expected reward. In a model with Mechanism 3, unequal reward information results in faster processing (i.e., earlier onset) of the target with higher expected reward. (C) Example behavior of a model with Mechanism 1. Probability of choosing the target on the more rewarding (better) side as a function of TOA in Experiment 1 (and $-TOA$ in Experiment 2) using Mechanism 1. The results with no reward modulation (black diamonds) are shown as the control. The model with Mechanism 1 (circles) produces similar shift in target selection in the two experiments. (D–E) Similar to C but for the model with Mechanisms 2 (D) and 3 (E). The models with reward-dependent modulation of visual input (Mechanisms 2 and 3) produce opposite shifts in the two experiments.

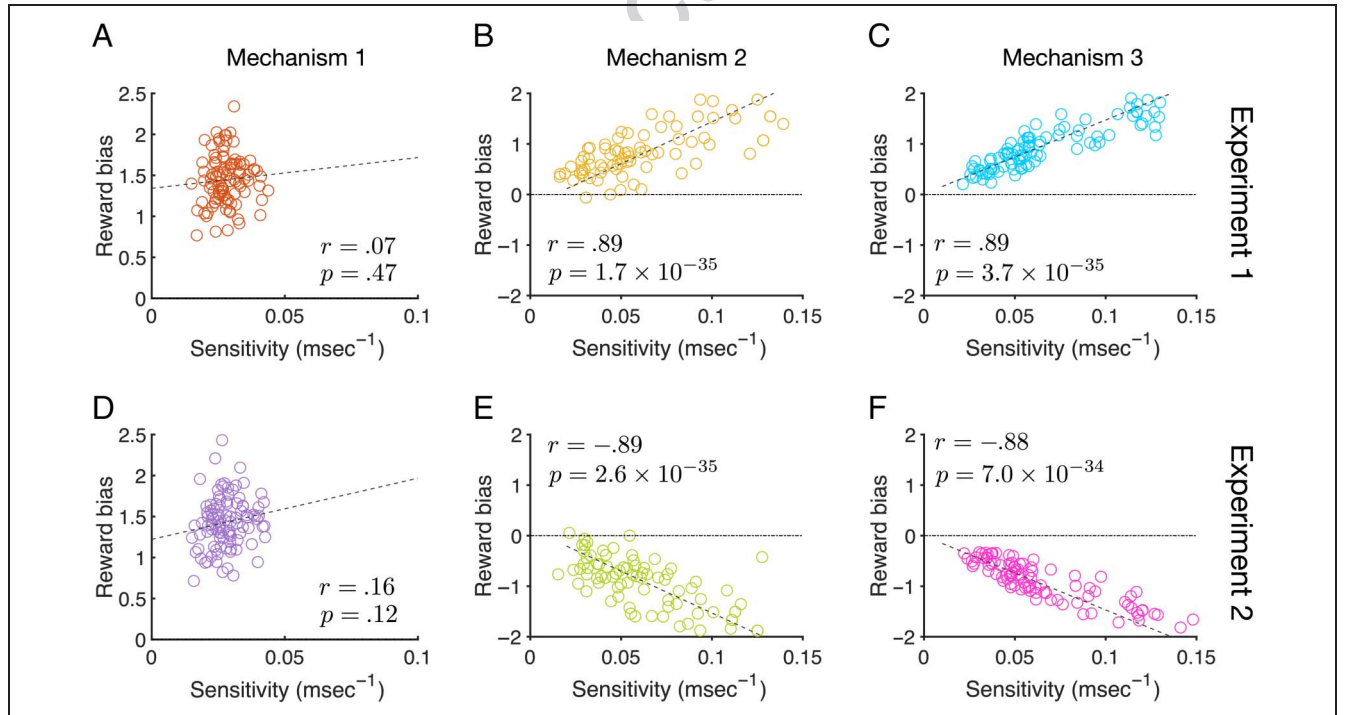


Figure 10. Shift in target selection due to unequal reward input was more compatible with Mechanism 1. (A) Reward bias is plotted as a function of the sensitivity to the TOA for Experiment 1 using the model based on Mechanism 1. Each point shows the replica of a participant. The gray dashed line shows the least squares line fitted on the simulated behavioral data. (B–C) The same as in A but using the model based on Mechanisms 2 (B) and 3 (C), respectively. (D–F) The same as in A–C but for Experiment 2. The model based on Mechanism 1 did not show a significant correlation between reward bias and the sensitivity to the TOA. However, the models based on Mechanisms 2 and 3 did exhibit significant correlation between reward bias and the sensitivity to the TOA.

assuming independent reward input (with additional noise) to later stage of decision-making, the model with Mechanism 1 can explain how reward biases could become independent of sensitivity to sensory signals and thus could not be optimized.

DISCUSSION

Several studies in the past two decades have aimed to reveal neural mechanisms by which reward influences perceptual decision-making. These studies have argued that unequal reward outcomes could either increase the tendency to choose the target with larger expected reward (response bias) and/or result in differential processing of sensory information and thus perceptual bias. To directly test these two alternative but not necessarily exclusive hypotheses, in two sets of experiments, we asked participants to saccade to the first or second target that appeared on the screen while we manipulated the amount of reward expected from the two alternative responses. Importantly, a bias in sensory processing would result in opposite shifts in target selection in the two experiments, whereas response bias would cause similar shifts. We did not find any evidence for different amounts of shift in the two experiments, indicating that expected reward is more likely to cause response bias rather than a bias in sensory processing. These findings dovetail with results from recent studies that used modeling to determine the mechanisms underlying the influence of expected reward on perceptual choice (Gao et al., 2011; Diederich, 2008; Diederich & Busemeyer, 2006) and studies that look at the effect of expectation in general (Rungratsameetaweemana, Itthipuripat, Salazar, & Serences, 2018; Bang & Rahnev, 2017).

Nonetheless, others have argued that reward can directly influence the processing of sensory information during perceptual decision-making (Cicmil et al., 2015; Liston & Stone, 2008; Pleger et al., 2008; Voss et al., 2008). A possible reason for the discrepancy between their findings and ours could be due to differences in the experimental paradigms in terms of time dependency. The temporal judgment task used here is a type of time-dependent perceptual choice, and it is possible that reward exerts its influence differently during time-independent perceptual choice. For example, the integration of sensory signal over time could push the influence of reward information to later stages of decision-making, resulting in response bias instead of perceptual bias. However, there are studies (e.g., Diederich & Busemeyer, 2006) showing the effects of reward as response bias even in discrimination between the lengths of two lines (i.e., time-independent tasks). Regardless, future studies are required (using our approach) to test the generalizability of our findings to other types of perceptual decision-making.

Similar to the effects of reward on perceptual decision-making, there is a long-lasting debate on whether attention influences perception by accelerating sensory

processing (the “prior entry” hypothesis) or inducing decision biases. Some have argued that attention enhances the speed of sensory processing (Hikosaka, Miyauchi, & Shimojo, 1993; Stelmach & Herdman, 1991), whereas others have maintained that observed effects are primarily due to attentional modifications of the decision mechanisms (Schneider & Bavelier, 2003). Most of these studies used a TOJ task to measure the point of subjective simultaneity (PSS) from attentional cueing. For example, Shore and colleagues examined the effect of attentional cueing on perception by asking participants to report the first or second targets that appeared on the screen to separate changes in sensory processing from response biases (Shore et al., 2001). They found that attention mainly influences perception by accelerating sensory processing. In contrast, Schneider and Bavelier (2003) have argued that the shift in the PSS due to attentional cueing in the TOJ task is not an adequate reason to accept the prior entry hypothesis. Instead, they suggest that one should compare shifts in the PSS in the TOJ task with those of in a simultaneity judgment task, in which the participants report whether two stimuli appeared simultaneously or successively. By making this comparison, they showed that attentional cueing has little influence on accelerating sensory processing similarly to what we found for the effects of unequal reward.

Here, rather than explicit attentional cueing, we used unequal reward information to bias processing of sensory information and/or decision-making, both of which could have behavioral benefits in terms of harvested reward. We also provided reward feedback (correct or incorrect judgment) on each trial, which allowed participants to correct their biases if desired so. The fact that we observed almost opposite results to those by Shore et al. (2001) based on attentional cueing indicates that reward information influences perception rather differently than how attention affects perception, and therefore, reward and attentional processes rely on different neural mechanisms to guide behavior. Furthermore, because reward information in our experiments was not predictive of the correct response, it is possible that this type of cueing exploits a different mechanism. Nonetheless, attention has been shown to closely interact with reward processing (Spitmaan, Chu, & Soltani, 2019; Farashahi, Azab, Hayden, & Soltani, 2018; Soltani, Khorsand, Guo, Farashahi, & Liu, 2016; Stănișor, van der Togt, Pennartz, & Roelfsema, 2013; Serences, 2008) and revealing that relationship is crucial for fully comprehending both processes (Maunsell, 2004).

Reward could also influence sensory processing without biasing sensory processing in a specific direction. For example, reward could cause arousal and/or increase motivation and effort in the task, both of which enhance sensory processing and performance as found in other studies (Vassena, Deraeve, & Alexander, 2019). Such enhancements would result in steeper psychometric function but

not a bias toward the better or worse option. However, we did not modulate the total expected reward nor did we change reward values between trials of a given condition (Neutral, Gain, or Loss). Thus, such motivational effects may have not been present or faded away quickly over the course of our experiments [R1.6].

It is important to note that, even though reward information was not predictive of the correct response in our experiments, participants still could use this information to obtain more reward. More specifically, increasing sensitivity to the more rewarding side does not help detection of the correct target but can improve performance in terms of obtained reward points because temporal judgment is not perfect. For example, an optimal observer such as one using the sequential ratio test can incorporate reward information to adjust decision criterion (Gold & Shadlen, 2001). Although an optimal observer may only change the threshold for response to the more rewarding side, a suboptimal observer may be persuaded to attend more to the more rewarding side (prior entry hypothesis), which could result in a change in perception. Interestingly, using visual search tasks in which different reward magnitudes were associated with detection of different objects, Hickey and colleagues have shown that reward enhances the saliency of certain features on subsequent trials, which results in suboptimal performance (Hickey, Chelazzi, & Theeuwes, 2010) [R1.6].

To reveal possible neural mechanisms underlying our observations, we extended a biophysically plausible cortical network model (Soltani et al., 2013) to simulate shifts in target selection due to unequal expected reward based on alternative mechanisms. We found that our experimental results are more compatible with the influence of reward information on later stages of decision-making processes via biasing the activity in the output layer of the decision circuit toward the target on the more rewarding side. Considering that lesions and reversible inactivation of the FEF cause similar shifts in target selection during the paired-target task (Schiller & Tehovnik, 2003; Schiller & Chou, 1998, 2000), reward could exert its influence through modulations of the output layer of the FEF. Furthermore, our analyses of SRT revealed that unequal expected reward outcomes resulted in faster decision-making. Future experiments that emphasize speed could provide additional information to test alternative models.

Importantly, we found that, to reproduce our experimental results by our model, the input for biasing target selection should be independent of sensory evidence, which is more consistent with results from a few recent studies (Gao et al., 2011; Diederich, 2008; Diederich & Busemeyer, 2006). For example, using a task in which the participants had to judge whether two lines are of the same or different length while manipulating the payoff for the two responses, Diederich and Busemeyer have shown that the effect of unequal payoffs is more compatible with a two-stage processing of sensory and reward information (Diederich & Busemeyer, 2006). In their

model, the decision maker first integrates reward information followed by the integration of sensory information (with no reward modulation) if no decision is made during the first stage. In another study, Gao and colleagues used a leaky competing accumulator model to show that reward information biases the initial state of the decision variable toward the target with higher expected reward (Gao et al., 2011). Both these studies illustrate that reward information does not interact with sensory evidence. In contrast, in our study, all the explored mechanisms generated comparable shifts in target selection in Experiment 1 suggests that, to distinguish the origin of reward effects, one needs to consider the appropriate task design in addition to the appropriate model.

Our results not only show that shifts in target selection due to unequal expected reward were suboptimal and independent of individuals' sensitivity to sensory signal but also explain that these shifts could not be optimized if reward influences later stages of decision-making independently of the sensory input. In addition, we observed a larger reward bias in the Loss condition compared with the Gain condition in both experiments. This result resembles loss aversion behavior during value-based choice (Tversky & Kahneman, 1992) and extends this phenomenon to perceptual decision-making with different reward outcomes. Together, these findings suggest that, even during perceptual choice, heuristics are used for differential processing of gain and loss information.

Similar but much weaker suboptimal behavior has also been observed for biased reward probabilities during perceptual decision-making (Navalpakkam, Koch, & Perona, 2009; Voss et al., 2008). Interestingly, it has been shown that humans exhibit a closer-to-optimal criterion when they deal with unequal reward probabilities rather than unequal reward magnitudes on alternative options or actions (Teichert & Ferrera, 2010; Maddox, 2002). Our modeling results indicate that shifts in target selection can be closer to optimal if reward information affects the processing of visual input instead of later stages of decision-making. Therefore, the difference in response to unequal reward probability and magnitude could be due to their influence on different stages of decision-making. Finally, in environments that resemble more naturalistic settings, adjustments in choice and learning of reward probability can occur in the absence of any optimization (Farashahi et al., 2017; Khorsand & Soltani, 2017). Future studies are required to determine whether reward probability and magnitude exert their influence at separate stages of decision-making.

Acknowledgments

We would like to thank Daeyeol Lee, Shiva Farashahi, and Mehran Spitmaan for their help in the early stages of this work and Patrick Cavanagh and Shih-Wei Wu for comments on the manuscript. This work was supported by the National Science Foundation under Grant 1632738 (A. S.).

Reprint requests should be sent to Alireza Soltani, Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, or via e-mail: soltani@dartmouth.edu.

REFERENCES

- Bang, J. W., & Rahnev, D. (2017). Stimulus expectation alters decision criterion but not sensory signal in perceptual decision making. *Scientific Reports*, 7, 17072.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10, 433–436.
- Carrasco, M., & Barbot, A. (2019). Spatial attention alters visual appearance. *Current Opinion in Psychology*, 29, 56–64.
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, 7, 308–313.
- Christopoulos, V., Bonaiuto, J., & Andersen, R. A. (2015). A biologically plausible computational theory for value integration and action selection in decisions with competing alternatives. *PLoS Computational Biology*, 11, e1004104.
- Christopoulos, V., & Schrater, P. R. (2015). Dynamic integration of value information into a common probability currency as a theory for flexible decision making. *PLoS Computational Biology*, 11, e1004402.
- Cicmil, N., Cumming, B. G., Parker, A. J., & Krug, K. (2015). Reward modulates the effect of visual cortical microstimulation on perceptual decisions. *eLife*, 4, e07832.
- Diederich, A. (2008). A further test of sequential-sampling models that account for payoff effects on response bias in perceptual decision tasks. *Perception & Psychophysics*, 70, 229–256.
- Diederich, A., & Busemeyer, J. R. (2006). Modeling the effects of payoff on response bias in a perceptual discrimination task: Bound-change, drift-rate-change, or two-stage-processing hypothesis. *Perception & Psychophysics*, 68, 194–207.
- Farashahi, S., Azab, H., Hayden, B., & Soltani, A. (2018). On the flexibility of basic risk attitudes in monkeys. *Journal of Neuroscience*, 38, 4383–4398.
- Farashahi, S., Donahue, C. H., Khorsand, P., Seo, H., Lee, D., & Soltani, A. (2017). Metaplasticity as a neural substrate for adaptive learning and choice under uncertainty. *Neuron*, 94, 401–414.
- Farashahi, S., Ting, C.-C., Kao, C.-H., Wu, S.-W., & Soltani, A. (2018). Dynamic combination of sensory and reward information under time pressure. *PLoS Computational Biology*, 14, e1006070.
- Feng, S., Holmes, P., Rorie, A., & Newsome, W. T. (2009). Can monkeys choose optimally when faced with noisy stimuli and unequal rewards? *PLoS Computational Biology*, 5, e1000284.
- Gao, J., Tortell, R., & McClelland, J. L. (2011). Dynamic integration of reward and stimulus information in perceptual decision-making. *PLoS One*, 6, e16749.
- Gold, J. I., & Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends in Cognitive Sciences*, 5, 10–16.
- Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience*, 30, 11096–11103.
- Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research*, 33, 1219–1240.
- Khorsand, P., & Soltani, A. (2017). Optimal structure of metaplasticity for adaptive learning. *PLoS Computational Biology*, 13, e1005630.
- Kleiner, M., Brainard, D. H., Pelli, D. G., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception*, 36, 1–16.
- Liston, D. B., & Stone, L. S. (2008). Effects of prior information and reward on oculomotor and perceptual choices. *Journal of Neuroscience*, 28, 13866–13875.
- Maddox, W. T. (2002). Toward a unified theory of decision criterion learning in perceptual categorization. *Journal of the Experimental Analysis of Behavior*, 78, 567–595.
- Maunsell, J. H. R. (2004). Neuronal representations of cognitive state: Reward or attention? *Trends in Cognitive Sciences*, 8, 261–265.
- Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience*, 32, 2335–2343.
- Navalpakkam, V., Koch, C., & Perona, P. (2009). Homo economicus in visual search. *Journal of Vision*, 9, 31.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Pleger, B., Blankenburg, F., Ruff, C. C., Driver, J., & Dolan, R. J. (2008). Reward facilitates tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *Journal of Neuroscience*, 28, 8161–8168.
- Rajsic, J., Perera, H., & Pratt, J. (2017). Learned value and object perception: Accelerated perception or biased decisions? *Attention, Perception, & Psychophysics*, 79, 603–613.
- Rorie, A. E., Gao, J., McClelland, J. L., & Newsome, W. T. (2010). Integration of sensory and reward information during perceptual decision-making in Lateral Intraparietal Cortex (LIP) of the macaque monkey. *PLoS One*, 5, e9308.
- Rungtameeetaweemana, N., Itthipuripat, S., Salazar, A., & Serences, J. T. (2018). Expectations do not alter early sensory processing during perceptual decision-making. *Journal of Neuroscience*, 38, 5632–5648.
- Schiller, P. H., & Chou, I.-H. (1998). The effects of frontal eye field and dorsomedial frontal cortex lesions on visually guided eye movements. *Nature Neuroscience*, 1, 248–253.
- Schiller, P. H., & Chou, I.-H. (2000). The effects of anterior arcuate and dorsomedial frontal cortex lesions on visually guided eye movements: 2. Paired and multiple targets. *Vision Research*, 40, 1627–1638.
- Schiller, P. H., & Tehovnik, E. J. (2003). Cortical inhibitory circuits in eye-movement generation. *European Journal of Neuroscience*, 18, 3127–3133.
- Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology*, 47, 333–366.
- Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron*, 60, 1169–1181.
- Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychological Science*, 12, 205–212.
- Soltani, A., Khorsand, P., Guo, C., Farashahi, S., & Liu, J. (2016). Neural substrates of cognitive biases during probabilistic inference. *Nature Communications*, 7, 11393.
- Soltani, A., Noudoost, B., & Moore, T. (2013). Dissociable dopaminergic control of saccadic target selection and its implications for reward modulation. *Proceedings of the National Academy of Sciences, U.S.A.*, 110, 3579–3584.
- Spitmaan, M., Chu, E., & Soltani, A. (2019). Salience-driven value construction for adaptive choice under risk. *Journal of Neuroscience*, 39, 5195–5209.
- Stanford, T. R., Shankar, S., Massoglia, D. P., Costello, M. G., & Salinas, E. (2010). Perceptual decision making in less than 30 milliseconds. *Nature Neuroscience*, 13, 379–385.
- Stănişor, L., van der Togt, C., Pennartz, C. M. A., & Roelfsema, P. R. (2013). A unified selection signal for attention and reward in primary visual cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, 110, 9136–9141.
- Stelmach, L. B., & Herdman, C. M. (1991). Directed attention and perception of temporal order. *Journal of Experimental*

- Psychology: Human Perception and Performance*, 17, 539–550.
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: Neural currencies for valuation and decision making. *Nature Reviews Neuroscience*, 6, 363–375.
- Summerfield, C., & Koechlin, E. (2010). Economic value biases uncertain perceptual choices in the parietal and prefrontal cortices. *Frontiers in Human Neuroscience*, 4, 208.
- Teichert, T., & Ferrera, V. P. (2010). Suboptimal integration of reward magnitude and prior reward likelihood in categorical decisions by monkeys. *Frontiers in Neuroscience*, 4, 186.
- Tosoni, A., Committeri, G., Calluso, C., & Galati, G. (2017). The effect of reward expectation on the time course of perceptual decisions. *European Journal of Neuroscience*, 45, 1152–1164.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.
- Vassena, E., Deraeve, J., & Alexander, W. H. (2019). Task-specific prioritization of reward and effort information: Novel insights from behavior and computational modeling. *Cognitive, Affective, & Behavioral Neuroscience*, 19, 619–636.
- Voss, A., Rothermund, K., & Brandtstädter, J. (2008). Interpreting ambiguous stimuli: Separating perceptual and judgmental biases. *Journal of Experimental Social Psychology*, 44, 1048–1056.
- Wong, K.-F., & Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *Journal of Neuroscience*, 26, 1314–1328.

Uncorrected Proof